SAND 2012-2496 P

Optimization-Based Modeling:

A New Strategy for the Compatible Discretization and Scalable Solution of Multiphysics Problems

Pavel Bochev

Numerical Analysis and Applications

Sandia National Laboratories

Exascale Research Conference, April 16-18, 2012, Portland

Supported in part by U.S. DEPARTMENT OF ENERGY | Office of Science

Sandia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000

SAND 2012-2496 P

Outline

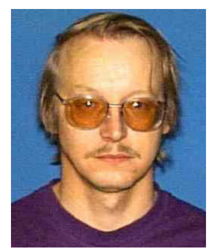
- **Research drivers**
- **Why optimization?**
- **Applications of Optimization-Based Modeling**
  - **Abstract theory of optimization based operator splitting**
  - **Application to synthesis of solvers**
  - **Robust and efficient optimization-based monotone transport**

Key Collaborators

K. Peterson
SNL

D. Ridzal
SNL

J. Young
SNL



M. Shashkov
LANL

M. Gunzburger
FSU

SAND 2012-2496 P

Research Drivers

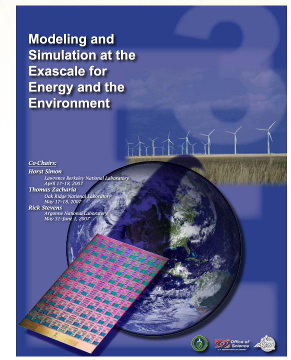
Robust and efficient solution of multiphysics problems

Dominant solution strategy for multiphysics, multiscale problems for 30+ years:

> 1^{st} order operator splitting, decoupled nonlinear solution methods, semi-implicit and explicit time integration.

This strategy is rapidly approaching a **point of diminishing returns** because

1) It lacks the stability properties for simulations over dynamic scales of interest
2) It often relies on heuristics to control the splitting errors
3) Is prone to non-intuitive instabilities



DOE Town Hall Report

Compatible discretization of multiphysics problems

The advanced state of **single physics discretizations** contrasts sharply with the **limited mathematical understanding** of compatible discretizations for multiphysics problems:

1) Lack of formal mathematical theory to guide the compatible discretization.
2) Physics components have disparate mathematical structures, which calls for mutually exclusive discretization and/or solver strategies.
3) Direct preservation of physical properties imposes severe grid/space constraints and tangles accuracy with the preservation of the properties.

Synthesis of Discretizations and Solvers

Challenges:

$$\partial_t u = (L_1 + L_2)u$$

Typically, L_1 and L_2 have different **mathematical structures**

$$\partial_t u^h = (L_1 + L_2)^h u^h$$

$$L_1^h \qquad L_2^h$$

Stable compatible methods may exist for L_1 and L_2 but not for the composite problem:

$$(L_1 + L_2)^h \neq L_1^h + L_2^h$$

$$\left(L_1^h\right)^{-1} \qquad \left(L_2^h\right)^{-1}$$

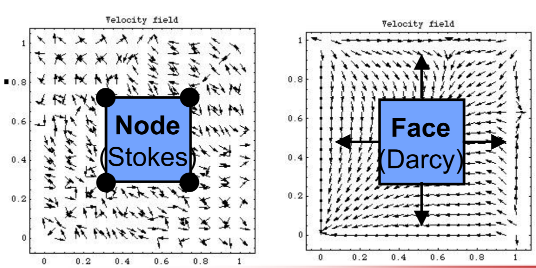
Efficient solvers may exist for L_1 and L_2 but not for the composite problem

$$\left((L_1 + L_2)^h\right)^{-1} = ?$$

Traditional approaches: regularization, operator splitting …

⇒ tunable parameters **reduce robustness**

⇒ splitting errors **reduce accuracy & stability**



Preservation of Physical Properties

Challenges:

$$\partial_t u = Lu$$

Generally, discretization does not automatically preserve constraints, even with stabilization/regularization

$$\partial_t u^h = L^h u^h$$

$$\underline{C} \le Cu \le \overline{C}$$

In multiphysics codes this solution is input for another physics component

$$\cancel{\underline{C} \le Cu^h \le \overline{C}}$$

$$Bu = b$$
$$\dots\dots\dots\dots\dots$$

Automatic preservation of maximum principle, local and global bounds, is required for robust, predictive simulations

$$\cancel{Bu^h = b}$$
$$\dots\dots\dots\dots\dots$$

Traditional approaches: limiters, “repair”, special grids, …

⇒ limiters & repair **entangle constraints & accuracy** and **obscure sources of discretization errors**

⇒ special grids **reduce the scope** of the methods

SAND 2012-2496 P

Optimization-based modeling (OBM)

Our approach: a divide and conquer strategy

Use **optimization and control ideas** to **manage externally** those objectives that are **difficult** (or impractical) to handle **directly** in the discretization process by manipulating the grid, the formulation, or the reconstruction.

Potential payoffs

- ➪ **Elimination of splitting errors**: reformulation into an equivalent optimization problem
- ➪ **Elimination of limiters**: lifts the associated restrictions on cell types & accuracy
- ➪ **Balancing of constraints**: accuracy, mass conservation, monotonicity, variable bounds…
- ➪ **Generality with respect to problem discretization**: applicable to FE, FV and FD schemes as well as particle methods, on mixed n-D grids
- ➪ **Generality with respect to problem type:** elliptic, hyperbolic, …
- ➪ **Enable efficient reuse of existing codes:** solvers, optimization tools,…

SAND 2012-2496 P

Synthesis of discretizations and solvers as an optimization problem

Objective	Constraints
Reconcile approximate solutions of the single physics operator equations	**Enforce** constituent component physics

$$\partial_t u = (L_1 + L_2)u$$

→

minimize $\frac{1}{2}\left(\left\|u_1^h - u_2^h\right\|^2 + \varepsilon\left\|\theta^h\right\|^2\right)$ ← **Reconcile** fields

subject to
$$\begin{cases} \gamma_1 \partial_t u_1^h = L_1^h u_1^h + \theta^h & \leftarrow \text{Physics operator \#1} \\ \gamma_2 \partial_t u_2^h = L_2^h u_2^h - \theta^h & \leftarrow \text{Physics operator \#2} \\ \gamma_1 + \gamma_2 = 1 & \leftarrow \text{Split ratio} \end{cases}$$

"Magic" Variable

References

- "*Optimization-based approach for robust solution algorithms*", Bochev, Ridzal, SINUM, 2009
- "*Additive operator decomposition*", Bochev, Ridzal, Springer LNCS 5910, 2010
- "*Optimization-based domain decomposition*", M. Gunzburger, 1997
- "*Decomposition of everything*", J.L. Lyons, 2001

SAND 2012-2496 P

Preservation of physical properties as an optimization problem

Objective	Constraints
Match a discrete target solution having the best possible accuracy	**Enforce** lost physical properties

$$\partial_t u = Lu$$
$$\underline{C} \le Cu \le \overline{C}$$
$$\ldots\ldots\ldots\ldots\ldots$$

$$\text{minimize} \quad \frac{1}{2}\left\| u^h - u_T^h \right\|^2$$

$$\text{subject to} \quad \begin{cases} \underline{C} \le Cu^h \le \overline{C} \\ \partial_t u_T^h = L^h u_T^h \end{cases}$$

Match target field

Enforce constraints

Define target field

References

- "*Optimization based remap*", Bochev, Ridzal, Scovazzi, Shashkov JCP, 2011
- "*Optimization-based transport*", Parts 1-3, Bochev, Peterson, Ridzal, Young, LNCS 2012

SAND 2012-2496 P

Abstract theory of additive operator splitting

Reformulation of $Q(u,v)=\langle f,v\rangle$ as a constrained optimization problem

min $J(u_1,u_2)=\frac{1}{2}\|u_1-u_2\|_U^2$ subject to $\begin{cases} Q_1(u_1,v_1)-(\theta,v_1)_V=\langle f,v_1\rangle & \forall v_1\in V \\ Q_2(u_2,v_2)+(\theta,v_2)_V=0 & \forall v_2\in V \end{cases}$

- u_1,u_2 – states
- θ – virtual control

Theorem

Assume that the additive split $Q(u,v)=Q_1(u,v)+Q_2(u,v)$ is such that

$$\sup_{v\in V}\frac{Q_i(u,v)}{\|v\|_V}\ge \underline{\gamma}_i\|u\|_U \qquad \sup_{u\in U}\frac{Q_i(u,v)}{\|u\|_U}>0 \quad \text{and} \quad Q_i(u,v)\le\bar{\gamma}_i\|u\|_U \quad \forall u\in U,\forall v\in V$$

There exist unique optimal solution (u_1,u_2,θ) and $u=u_1=u_2$ where $Q(u,v)=\langle f,v\rangle$.

Notable facts

- Optimization **exposes** the **constituent components** of the **multiphysics** operator
- Optimization problem is **well-posed without control penalty**
- As a result, **original** and **reformulated** problems are **completely equivalent**

There's no splitting error!

SAND 2012-2496 P

Application: Synthesis of Fast Solvers

Assumptions

$$Q(u,v) = Q_1(u,v) + Q_2(u,v)$$

Fast and efficient solvers exist for Q_1 and Q_2

Approach: solve the equivalent *reduced-space* optimization problem

$$\min J(\vec{u}_1,\vec{u}_2) = \frac{1}{2}(\vec{u}_1 - \vec{u}_2)^T \mathbf{U}(\vec{u}_1 - \vec{u}_2) \text{ s.t. } \begin{cases} \mathbf{Q}_1\vec{u}_1 - \mathbf{V}\vec{\theta} = \vec{f} \\ \mathbf{Q}_2\vec{u}_2 + \mathbf{V}\vec{\theta} = \vec{0} \end{cases} \Rightarrow \min J_{RED}(\vec{\theta}) = \frac{1}{2}\vec{\theta}^T \mathbf{H}_{RED}\vec{\theta} - \vec{\theta}^T \vec{f}_{RED}$$

Algorithm:

Equation	Compute	Solve	Properties
Adjoint	$\vec{y}_1 = \mathbf{V}\vec{\theta}$	$\mathbf{Q}_1\vec{x}_1 = \vec{y}_1, \quad \mathbf{Q}_2\vec{x}_2 = \vec{y}_1$	**Concurrency:** state and adjoint can be solved independently.
State	$\vec{y}_2 = \mathbf{U}(\vec{x}_1 - \vec{x}_2)$	$\mathbf{Q}_1^T\vec{x}_3 = \vec{y}_2, \quad \mathbf{Q}_2^T\vec{x}_4 = -\vec{y}_2$	**Efficiency:** application of H_{RED} only requires inversion of operators for which fast solvers exist.
$\mathbf{H}_{RED}\vec{\theta} = \vec{f}_{RED}$	$\mathbf{H}_{RED}\vec{\theta} = \mathbf{V}(\vec{x}_3 + \vec{x}_4)$		

Sandia National Laboratories

SAND 2012-2496 P

Application to an advection-diffusion problem

Additive split γ = 1

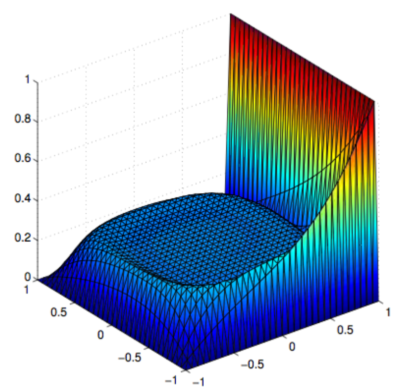
$Q_1(u,v)$ = Diffusion	$Q_2(u,v)$ = -Diffusion
$\gamma\left(\nabla u^h,\nabla v^h\right)+\left(\mathbf{b}\cdot\nabla u^h, v^h+\tau\mathbf{b}\cdot\nabla v^h\right)$	$(\kappa-\gamma)\left(\nabla u^h,\nabla v^h\right)-\left\langle \kappa\Delta u^h, \tau\mathbf{b}\cdot\nabla v^h\right\rangle_h$

Synthesized solver

GMRES(200)	ML^SGS	ML^SGS
$\mathbf{H}_{RED}\vec{\theta} = \vec{f}_{RED}$	$\mathbf{Q}_1 \quad \mathbf{Q}_1^T$	$\mathbf{Q}_2 \quad \mathbf{Q}_2^T$

Elman/Silvester/Wathen: "double-glazing" $\mathbf{b} \neq const$

Essentially fixed cost



Study	Fixed diffusion: 10^{-8}			Fixed grid size: 128		
Solver ↓	64	128	256	10^{-2}	10^{-4}	10^{-8}
Synthesized	**114**	**97**	**77**	**62**	**97**	**97**
ML^SGS	**97**	**ST**	**ST**	**11**	**ST**	**ST**
ML^ILU	71	196	MX	9	96	196
BAMG	72	457	MX	7	33	457

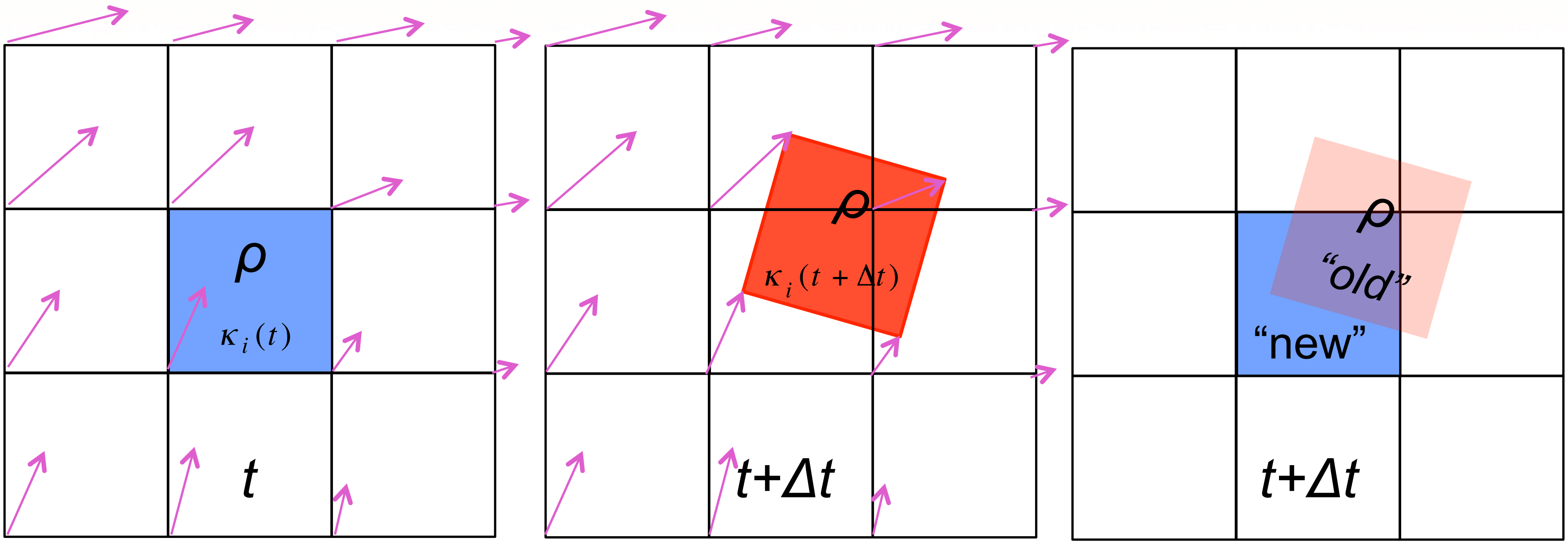
BAMG= Boomer AMG from hypre (LLNL) **ML = Trilinos AMG (Sandia)**

"Optimization-based approach for robust solution algorithms", Bochev, Ridzal, SINUM, 2009

Sandia National Laboratories

Optimization-based monotone transport (OBT)

Mass is conserved in Lagrangian volumes: $\frac{d}{dt} m_i(t) = \frac{d}{dt} \int_{\kappa_i(t)} \rho(x) dV = 0$



$m_i(t) = \int_{\kappa_i(t)} \rho(x) dV$

$m_i(t+\Delta t) = m_i(t)$

Remap

Transport = incremental mass/density remap

Bochev, Ridzal, Young, Peterson. *Optimization-based modeling with applications to transport Parts 1-3*, **Springer Lecture Notes in Computer Science**, *LNCS 7116, 2012.*

SAND 2012-2496 P

Mass/density remap as an optimization problem

The exact mass on new cell $\tilde{\kappa}_i$ can be expressed in aggregate mass-transfer form:

$$\tilde{m}_i^{EX} = m_i^{EX} + \delta m_i^{EX}; \quad \delta m_i^{EX} = \int_{\tilde{\kappa}_i} \rho(x)dV - \int_{\kappa_i} \rho(x)dV$$

Therefore, the mass on the new cell $\tilde{\kappa}_i$ can be approximated by

$$\tilde{m}_i^h = m_i^h + \delta m_i^h, \qquad \text{where} \qquad \delta m_i^h = \int_{\tilde{\kappa}_i} \rho_i^h(x)dV - \int_{\kappa_i} \rho_i^h(x)dV \approx \delta m_i^{EX}$$

C1: Mass conservation. Requires a single linear constraint:

$$\sum_{Cell} \delta m_i^h = 0 \quad \Rightarrow \quad \sum_{Cell} \tilde{m}_i^h = M$$

C2: Linearity preservation. Guaranteed if ρ_i^h is exact for linear functions on all κ_i:

$$\delta m_i^T = \int_{\tilde{\kappa}_i} \rho_i^h(x)dV - \int_{\kappa_i} \rho_i^h(x)dV$$ → Target (**high-order**) mass-transfers

C3: Local bounds ⇒ $\delta \tilde{m}_i^{min} \le \sum_{cell} \delta m_i^h \le \delta \tilde{m}_i^{max} \quad i = 1,\ldots,N$ → **Box constraints**

Mass/density remap as a QP

OBT = "*singly linearly constrained QP with simple bounds*"

$$\underset{\delta m_i^h}{\text{minimize}} \quad \sum_{Cell} \left(\delta m_i^h - \delta m_i^T\right)^2 \quad \text{subject to} \qquad \Leftarrow \textbf{C2}$$

$$\delta \tilde{m}_i^{min} \le \delta m_i^h \le \delta \tilde{m}_i^{max} \quad i = 1,\ldots,N \qquad \Leftarrow \textbf{C3}$$

$$\sum_{Cell} \delta m_i^h = 0 \qquad \Leftarrow \textbf{C1}$$

Theorem.

Existence of unique optimal solutions.
The OBT feasible set is non-empty: given a density distribution there exists a set of aggregate mass transfers δm_i^h, which satisfy the box constraints and sum up to zero.

Preservation of linearity.
Under mild conditions on the mesh motion, OBT preserves linear densities.

SAND 2012-2496 P

Fast Optimization Algorithm for OBT

Key property of singly linearly constrained QP with simple bounds:

$$\underset{\delta m_i^h}{\text{minimize}} \quad \sum_{Cell} \left(\delta m_i^h - \delta m_i^T\right)^2 \quad \text{subject to}$$

$$\delta \tilde{m}_i^{min} \le \delta m_i^h \le \delta \tilde{m}_i^{max}; \; i = 1,\ldots,N \quad \text{and} \quad \sum_{Cell} \delta m_i^h = 0$$

Without the equality constraint the QP is fully separable into *N* one-dimensional QPs with simple bounds

The Lagrangian

$$L(\delta m, \lambda, \mu_1, \mu_2) = \sum_{Cell} \left(\delta m_i^h - \delta m_i^T\right)^2 - \lambda \sum_{Cell} \delta m_i^h - \sum_{Cell} \mu_{1,i}(\delta m_i^h - \delta \tilde{m}_i^{min}) - \sum_{Cell} \mu_{2,i}(\delta m_i^h - \delta \tilde{m}_i^{max})$$

The Karush-Kuhn-Tucker (KKT) conditions

$$\begin{cases} \delta m_i^h = \delta m_i^T + \lambda + \mu_{1,i} - \mu_{2,i} \\ \delta \tilde{m}_i^{min} \le \delta m_i^h \le \delta \tilde{m}_i^{max} \\ \mu_{1,i} \ge 0, \quad \mu_{2,i} \ge 0 \\ \mu_{1,i}(\delta m_i^h - \delta \tilde{m}_i^{min}) = 0, \\ \mu_{2,i}(\delta m_i^h - \delta \tilde{m}_i^{max}) = 0 \end{cases} \quad \text{and} \quad \sum_{Cell} \delta m_i^h = 0$$

Without the equality constraint the KKT conditions are fully separable and can be solved in parallel for any fixed value of *λ*.

Fast Optimization Algorithm for OBT

Step 1: solve for *λ* fixed

$$\delta m_i^h = \delta m_i^T + \lambda \qquad \mu_{1,i} = 0 \qquad \mu_{2,i} = 0 \qquad \text{if} \quad \delta\tilde{m}_i^{min} \le \delta m_i^T + \lambda \le \delta\tilde{m}_i^{max}$$

$$\delta m_i^h = \delta\tilde{m}_i^{min} \qquad \mu_{2,i} = 0 \qquad \mu_{1,i} = \delta m_i^h - \delta m_i^T - \lambda \qquad \text{if} \quad \delta\tilde{m}_i^{min} \ge \delta m_i^T + \lambda$$

$$\delta m_i^h = \delta\tilde{m}_i^{max} \qquad \mu_{1,i} = 0 \qquad \mu_{2,i} = \delta m_i^T - \delta m_i^h + \lambda \qquad \text{if} \quad \delta m_i^T + \lambda \ge \delta\tilde{m}_i^{max}$$

$$\Rightarrow \quad \delta m_i^h(\lambda) = median(\delta\tilde{m}_i^{min}, \delta m_i^T + \lambda, \delta\tilde{m}_i^{max}); \quad i = 1,....,N$$

Step 2: adjust *λ* in an outer iteration to satisfy the single equality constraint

Solve $\sum_{Cell} \delta m_i^h(\lambda) = 0$ $\Rightarrow$ **piecewise linear, monotonically increasing function of single scalar variable *λ*.**

- Can solve to machine precision by a simple secant method
- Globalization is unnecessary because $\lambda_0=0$ is an excellent initial guess:

$$\delta m_i^h(\lambda_0) = median(\delta\tilde{m}_i^{min}, \delta m_i^T, \delta\tilde{m}_i^{max}); \quad i = 1,....,N$$

- $\delta m_i^h(\lambda_0)$ solves the QP without the equality constraint, i.e., "almost" a solution
- Locality $\Rightarrow \delta m_i^h(\lambda_0)$ barely violates the mass conservation constraint

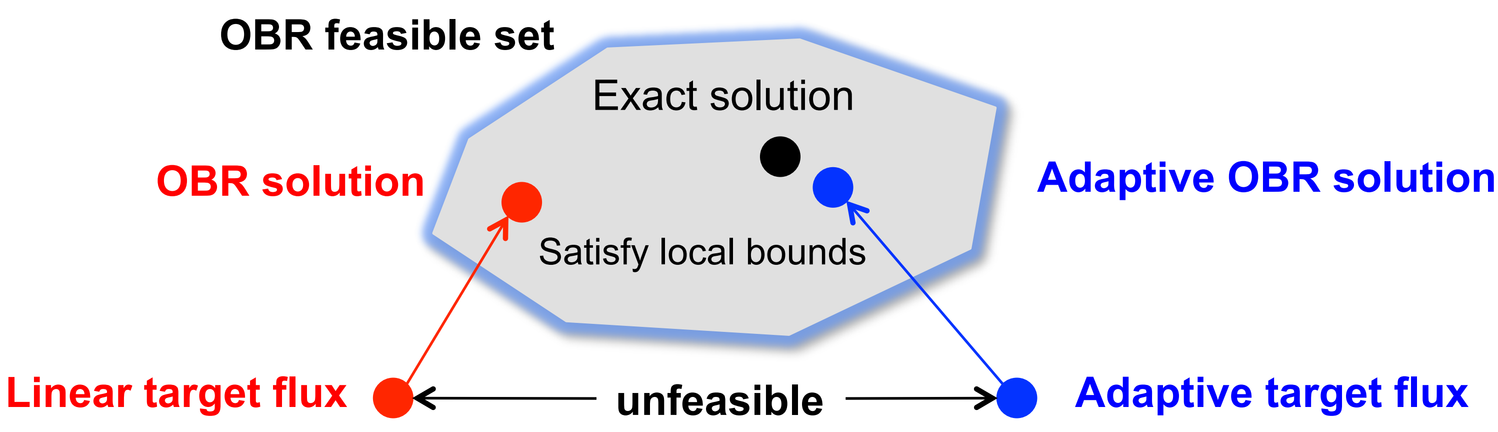
OBT with adaptive targets

OBT always finds the best possible (optimal) solution w.r.t. the targets

- We can improve OBR/T solution by using **targets, which adapt to local** solution features

Adaptive target definition

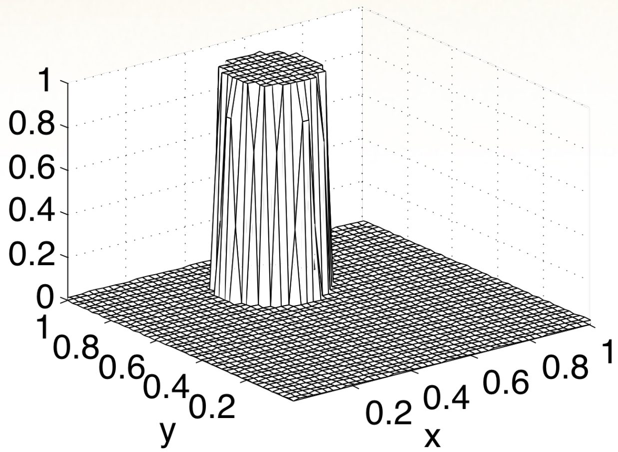
- **Use residual information to modify targets depending on local solution features**



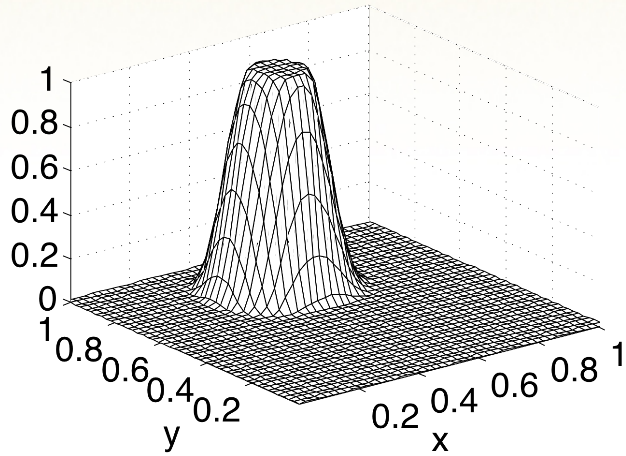
Because OBR/OBT completely separates reconstruction and bounds enforcement,

Target fluxes can **adapt to problem features** without concern for the bounds – the QP constraints will take care to enforce the bounds later!

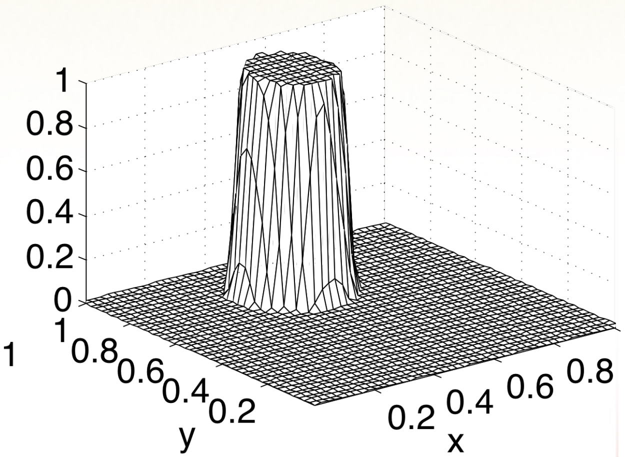
OBT with adaptive targets: cylinder



Initial



OBT



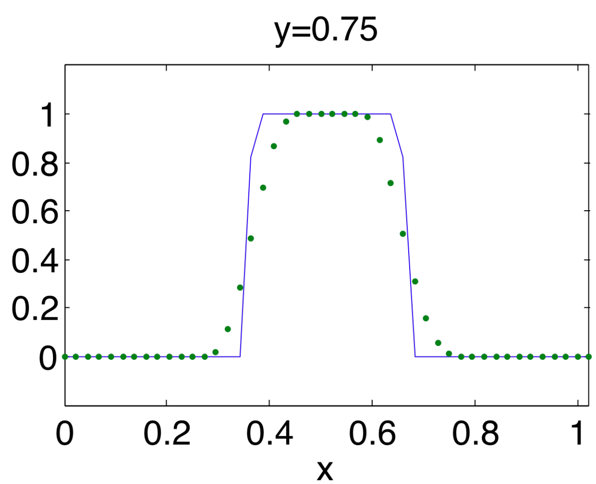
OBT-A

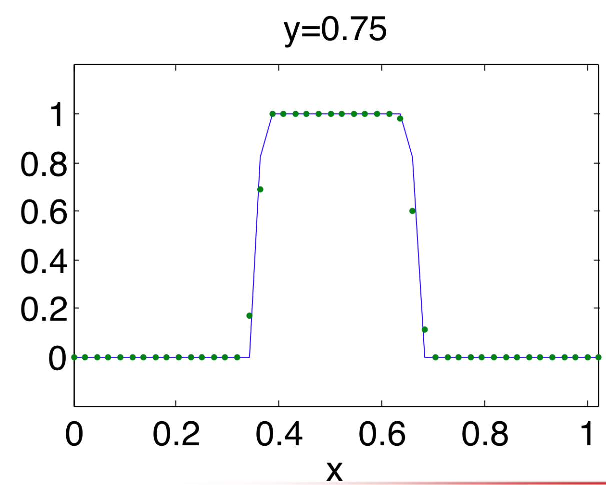
Rotating cylinder

$$u = -(y - 0.5) \qquad v = (x - 0.5)$$

Grid size: *N*x*N*, ***N*=45**

Time steps: $2\pi N$ **282**





OBT with adaptive targets: combo

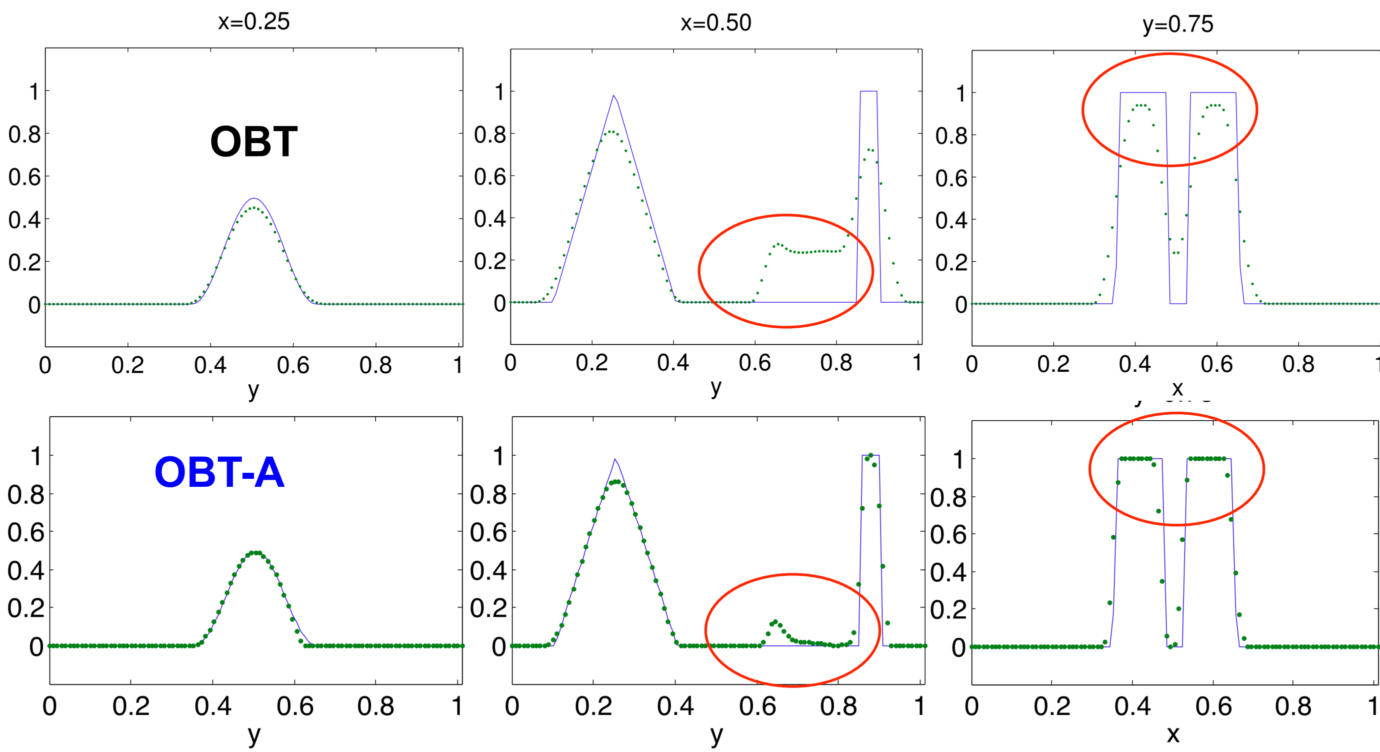
Rotating flow example (LeVeque, SINUM 33, 1996)

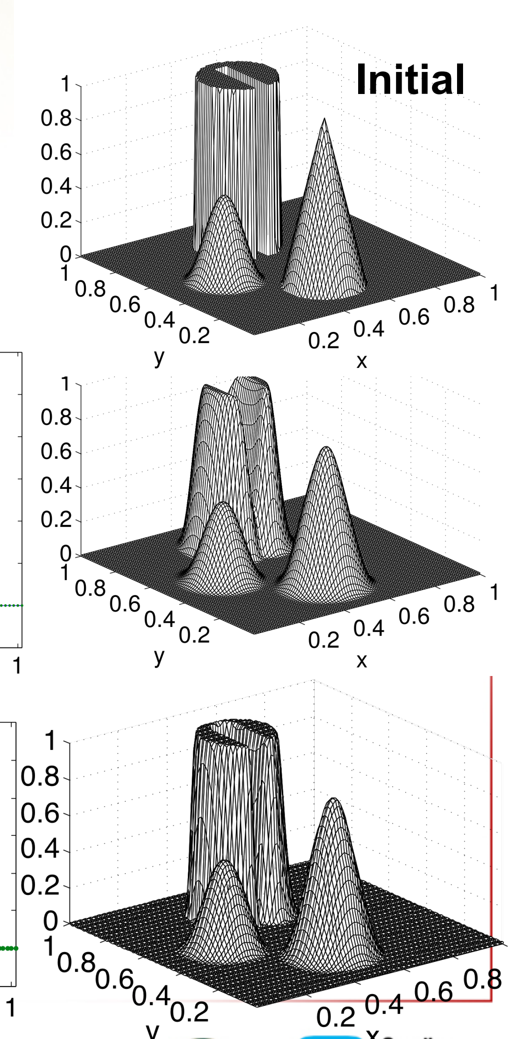
$u = -(y - 0.5) \qquad v = (x - 0.5)$

Grid size: *N*x*N*, ***N*=100**

Time steps: $2\pi N$ **628**

CFL < 1





This example combines "**smooth**" and "**sharp**" features!

OBT is as efficient as explicit transport

Matlab **wall-clock times on a 3.06GHz Intel Core Duo MacBook Pro**

"Cone"

Cells	Time steps	FCR (sec)	Van Leer	OBT	OBT/FCR
64x64	400	4.59	4.50	**4.92**	**1.1**
128x128	810	44.64	47.25	**48.62**	**1.1**
256x256	1,610	387.88	393.64	**403.23**	**1.0**
512x512	3,220	5,715.08	5,804.66	**5655.06**	**0.9**

"Combo"

Cells	Time steps	FCR (sec)	Van Leer	OBT	OBT/FCR
64x64	400	4.51	4.55	**4.98**	**1.1**
128x128	810	47.60	48.35	**48.78**	**1.0**
256x256	1,610	390.47	399.15	**405.92**	**1.0**
512x512	3,220	5802.05	5804.66	**5,655.11**	**0.9**

FCR = Flux Corrected Remap, Liska, et al, JCP 2010

SAND 2012-2496 P

Yet, OBT has superior robustness and accuracy

Preservation of monotonicity

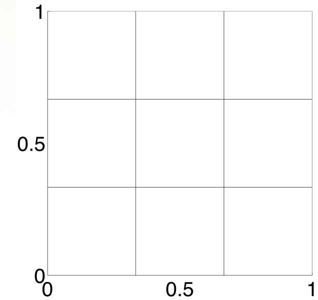
	C=5	C=6	C=7	C=14	C=15	C=16	C=100
OBT	✔	✔	✔	✔	✔	✔	✔
FCR	✔	✔	✔	✗	✗	✗	✗

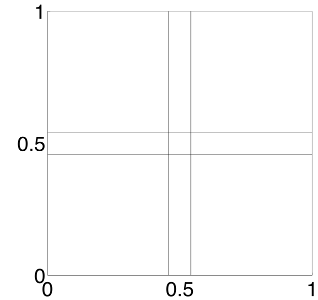
Preservation of linearity

	C=3	C=4	C=5	C=15	C=16	C=100
OBT	✔	✔	✔	✔	✔	✔
FCR	✔	✗	✗	✗	✗	✗

Rates of convergence

Sine & repeated repair		***FCR***		***OBT***	
#Cells	#remaps	L_1 error	L_1 rate	L_1 error	L_1 rate
128x128	640	2.81E-04	-	2.77E-04	-
256x256	1280	9.23E-05	**1.61**	6.82E-05	**2.04**
512x512	2560	3.65E-05	**1.47**	1.69E-05	**2.03**
1024x1024	5120	1.69E-05	**1.35**	4.18E-06	**2.00**





Mesh motion:
“Repeated repair”

Summary

A divide and conquer strategy: we use optimization ideas to separate discretization from tasks that are difficult to accomplish directly

Abstract theory for optimization-based additive operator splitting

- Increases concurrency by exposing constituent physics components
- Remove order & stability limitations (**no splitting error**)
- Rigorous mathematical foundations inherited from rich optimization theory
- Enables reuse of software components through synthesis of solvers and discretizations – compatible with the PETSc strategy for "*composable extreme-scale solvers*"

Optimization-based conservative and monotone transport (OBT)

- Completely separates accuracy from the enforcement of bounds:
  - ✓ sources of error traceable!
  - ✓ targets can be adapted to local solution features
- OBT is global QP: yields the best possible, w.r.t. the objective, solution
  - ✓ Increases robustness: can run at higher CFL numbers
  - ✓ Increases accuracy: remains 2nd order under most challenging mesh motions
- Yet, resulting QP can be solved efficiently: cost = cost of explicit methods